CLAIMS

This listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

1. (Currently Amended) A method for discovering knowledge from a set of text documents using a processor, the method comprising ~~the steps of:~~

extracting ~~from text documents~~ semi-structured meta-data from the set of text documents using a meta-data extractor, ~~wherein~~ the semi-structured meta-data comprising ~~includes~~ a plurality of concepts ~~entities~~ and a plurality of relations between the concepts ~~entities~~;

filtering the semi-structured meta-data to identify~~ing from the semi-structured meta-data~~ a ~~plurality~~ set of key concepts ~~entities~~ and a corresponding ~~plurality~~ set of key relations between the key concepts, the set of key concepts corresponding to the plurality of concepts;

~~deriving from a domain knowledge base a plurality of attributes relating to each of the plurality of entities relating to one of the plurality of key entities for forming a plurality of pairs of key entity and a plurality of attributes related thereto;~~

deriving at least one set of sub-concepts corresponding to the set of key concepts based upon data within a domain knowledge base, using a meta-data transformer;

formulating a plurality of training samples ~~patterns~~, each training sample including a vector representing a sub-concept and a vector representing a key concept ~~of the plurality of patterns relating to one of the plurality of pairs of key entity and a plurality of attributes related thereto~~; and

analyzing the plurality of training samples ~~patterns~~ using an associative discoverer to derive a set of associations between a set of vectors representing a sub-concept and at least one vector representing a key concept.~~; and~~

~~interpreting the output of the associative discoverer for discovering knowledge.~~

wherein neither the set of text documents nor the semi-structured meta-data mention the set of associations, and

wherein the set of associations corresponds to discovered knowledge that is extractable by a knowledge interpreter.

2. (Currently Amended) The method as in claim 1, wherein ~~the step of~~ extracting semi-structured meta-data from the set of text documents ~~from text documents~~ comprises ~~the step of~~ extracting text content from documents containing at least one type of text, image, audio, and video information.

3. (Currently Amended) The method as in claim 1, wherein ~~the step of identifying the plurality of key entities~~ filtering the semi-structured meta-data comprises ~~the step of~~ selecting the set ~~plurality~~ of key concepts ~~entities~~ according to frequency of appearance of the ~~plurality~~ set of key concepts ~~entities~~ in the semi-structured meta-data.

4. (Currently Amended) The method as in claim 1, wherein ~~the step of~~ identifying the set ~~plurality~~ of key relations comprises ~~the step of~~ selecting the set ~~plurality~~ of key relations according to frequency of appearance of the set ~~plurality~~ of key relations in the semi-structured meta-data.

5. (Currently Amended) The method as in claim 1, wherein ~~the step of deriving from the domain knowledge base comprises the step of deriving from a~~ data within the domain knowledge base relate ~~relating~~ to at least one of taxonomy, a concept hierarchy network, ontology, a thesaurus, a relational database, and an object-oriented database.

6. (Cancelled)

7. (Currently Amended) The method as in claim 1, wherein ~~step of the~~ formulating the plurality of ~~patterns~~ training samples comprises ~~the step of~~ formulating concatenated vector representations of the ~~plurality of attributes~~ set of sub-concepts and the set ~~plurality~~ of key concepts ~~entities~~ relating to the corresponding ~~plurality~~ set of key relations.

8. (Currently Amended) The method as in claim 1, wherein ~~the step of~~ analyzing the plurality of training samples ~~patterns~~ using the associative discoverer comprises the step of analyzing the plurality of training samples ~~patterns~~ using at least one of a neural network, a statistical system, and a symbolic machine learning system.

9. (Currently Amended) The method as in claim 8, wherein ~~the step of~~ analyzing the plurality of training samples ~~patterns~~ comprises the step of analyzing the plurality of training samples ~~patterns~~ using an Adaptive Resonance Associative Map.

10. (Currently Amended) The method as in claim 1, wherein ~~the step of interpreting the output of the associative discoverer for discovering knowledge~~ extraction of discovered knowledge by the knowledge interpreter comprises ~~the step of~~ discovering the semantic relations between the set ~~plurality~~ of ~~attributes~~ sub-concepts and the set ~~plurality~~ of key concepts ~~entities~~.

11. (Currently Amended) The method as in claim 1, further comprising the step of using a user interface for displaying the semi-structured meta-data, the set ~~plurality~~ of key concepts ~~entities~~, the set ~~plurality~~ of key relations, the set ~~plurality~~ of sub-concepts ~~attributes~~, and the knowledge discovered.

12. (Currently Amended) The method as in claim 1, further comprising the step of using a user interface for obtaining user instruction for the set ~~plurality~~ of key concepts ~~entities~~ and the set ~~plurality~~ of key relations.

13. (Currently Amended) A computer program product comprising a computer usable medium having computer readable program code means embodied in the medium for discovering knowledge from a set of text documents, the computer program product comprising:

computer readable program code means for extracting semi-structured meta-data from the set of text documents using a meta-data extractor, the semi-structured meta-data comprising a plurality of concepts and a plurality of relations between the concepts ~~extracting from text documents semi-structured meta-data, wherein the semi-structured meta-data includes a plurality of entities and a plurality of relations between the entities~~;

computer readable program code means for filtering the semi-structured meta-data to identify a set of key concepts and a corresponding set of key relations between the key concepts, the set of key concepts corresponding to the plurality of concepts ~~identifying from the semi-structured meta-data a plurality of key entities and a corresponding plurality of key relations~~;

computer readable program code means for deriving at least one set of sub-concepts corresponding to the set of key concepts based upon data within a domain knowledge base, using a meta-data transformer ~~deriving from a domain knowledge base a plurality of attributes relating to each of the plurality of entities relating to one of the plurality of key entities for forming a plurality of pairs of key entity and a plurality of attributes related thereto~~;

computer readable program code means for formulating a plurality of training samples, each training sample including a vector representing a sub-concept and a vector representing a key concept ~~formulating a plurality of patterns, each of the plurality of patterns relating to one of the plurality of pairs of key entity and a plurality of attributes related thereto~~; and

computer readable program code means for analyzing the plurality of training samples using an associative discoverer to derive a set of associations between a set of vectors representing a sub-concept and at least one vector representing a key concept, ~~analyzing the plurality of patterns using an associative discoverer; and~~

~~computer readable program code means for interpreting the output of the associative discoverer for discovering knowledge.~~

wherein neither the set of text documents nor the semi-structured meta-data mention the set of associations, and

wherein the set of associations corresponds to discovered knowledge that is extractable by a knowledge interpreter.

14. (Currently amended) The computer program product as in claim 13, wherein the computer readable program code means for extracting semi-structured meta-data from the set of text documents comprises ~~extracting from text documents comprises~~ computer readable program code means for extracting text content from documents containing at least one of text, image, audio, and video information.

15. (Currently Amended) The computer program product as in claim 13, wherein the computer readable program code means for filtering the semi-structured meta-data ~~identifying the plurality of key entities~~ comprises computer readable program code means for selecting the set of key concepts according to frequency of appearance of the set of key concepts in the semi-structured meta-data. ~~selecting the plurality of key entities according to frequency of appearance of the plurality of key entities in the semi-structured meta-data.~~

16. (Currently Amended) The computer program product as in claim 13, wherein the computer readable program code means for identifying the set ~~plurality~~ of key relations comprises computer readable program code means for selecting the set of key relations according to frequency of appearance of the set of key relations in the semi-structured meta-data. ~~selecting the plurality of key relations according to frequency of appearance of the plurality of key relations in the semi-structured meta-data.~~

17. (Currently Amended) The computer program product as in claim 13, wherein ~~the computer readable program code means for deriving from the domain knowledge base comprises computer readable program code means for deriving from a domain knowledge base relating~~ the data within the domain knowledge base relate to at least one of taxonomy, a concept hierarchy network, ontology, a thesaurus, a relational database, and an object-oriented database.

18. (Cancelled)

19. (Currently Amended) The computer program product as in claim 13, wherein the computer readable program code means for formulating the plurality of training samples ~~patterns~~ comprises computer readable program code means for formulating concatenated vector representations of the set of sub-concepts and the set of key concepts relating to the corresponding set of key relations. ~~formulating concatenated vector representations of the plurality of attributes and the plurality of key entities relating to the corresponding plurality of key relations.~~

20. (Currently Amended) The computer program product as in claim 13, wherein the computer readable program code means for analyzing the plurality of training samples ~~analyzing the plurality of patterns~~ using the associative discoverer comprises computer readable program code means for analyzing the plurality of training samples ~~analyzing the plurality of patterns~~ using at least one of a neural network, a statistical system, and a symbolic machine learning system.

21. (Currently Amended) The computer program product as in claim 20, wherein the computer readable program code means for analyzing the plurality of training samples ~~patterns~~ comprises computer readable program code means for analyzing the plurality of training samples ~~patterns~~ using an Adaptive Resonance Associative Map.

22. (Currently Amended) The computer program product as in claim 13, wherein extraction of discovered knowledge by the knowledge interpreter comprises ~~the computer readable program code means for interpreting the output of the associative discoverer for discovering knowledge comprises computer readable program code means for~~ discovering the semantic relations between ~~the plurality of attributes and the plurality of key entities.~~ the set of sub-concepts and the set of key concepts.

23. (Currently Amended) The computer program product as in claim 13, further comprising computer readable program code means for using a user interface for displaying the semi-structured meta-data, the set of key concepts, the set of key relations, the set of sub-concepts ~~the plurality of key entities, the plurality of key relations, the plurality of attributes~~, and the knowledge discovered.

24. (Currently Amended) The computer program product as in claim 13, further comprising computer readable program code means for using a user interface for obtaining user instruction for ~~the plurality of key entities and the plurality of key relations.~~ the set of key concepts and the set of key relations.

25. (Currently Amended) A system for knowledge discovery from a set of free-text documents, the system comprising:

means for extracting semi-structured meta-data from the set of free-text documents;

means for ~~identifying key entities and key relations from the semi-structured meta-data~~ filtering the semi-structured meta-data to identify a set of key concepts and a corresponding set of key relations between the key concepts, the set of key concepts corresponding to the plurality of concepts;

~~a knowledge base that defines the attributes of entities;~~

means for deriving at least one set of sub-concepts corresponding to the set of key concepts based upon data within a domain knowledge base,

means for formulating training samples, each training sample including a vector representing a sub-concept and a vector representing a key concept ~~patterns based on the key entities and the attributes of entities related to the key entities;~~ and

means for deriving a set of associations between a set of vectors representing a sub-concept and at least one vector representing a key concept, ~~analyzing the patterns for knowledge~~.

wherein neither the set of text documents nor the semi-structured meta-data mention the set of associations, and

wherein the set of associations corresponds to discovered knowledge that is extractable by a knowledge interpreter.

26. (Currently Amended) The system according to claim 25 wherein the semi-structured meta-data comprises definition of concepts ~~entities~~ and semantic relations among the concepts ~~entities~~.

27. (Previously Amended) The system according to claim 25 wherein the semi-structured meta-data is stored in at least one of a permanent and temporary storage.

28. (Currently Amended) The system according to claim 25 wherein the set of free-text documents comprise text, image, audio, video, or any combination thereof.

29. (Currently Amended) The system according to claim 25 wherein the means for ~~identifying key entities selects entities according to the key entities' frequency of appearance in the semi-structured meta-data.~~ filtering the semi-structured meta-data comprises means for selecting the set of key concepts according to frequency of appearance of the set of key concepts in the semi-structured meta-data.

30. (Currently Amended) The system according to claim 25 wherein the means for identifying ~~key relations selects relations according to the key relations' frequency of appearance in the semi-structured meta-data.~~ the set of key relations comprises means for selecting the set of key relations according to frequency of appearance of the set of key relations in the semi-structured meta-data.

31. (Currently Amended) The system according to claim 25 wherein the domain knowledge base comprises a taxonomy, a concept hierarchy network, an ontology, a thesaurus, a relational database, an object-oriented database, or any combination thereof.

32. (Cancelled)

33. (Currently Amended) The system according to claim 25 wherein the training samples ~~examples~~ comprises ~~concatenated vectors of the key entities, and the attributes of entities related to the key entities with a key relation.~~ concatenated vector representations of the set of sub-concepts and the set of key concepts relating to the corresponding set of key relations.

34. (Currently Amended) The system according to claim 25 wherein the means for ~~analyzing the patterns for knowledge~~ deriving the set of associations comprises a neural network, a statistical system, a symbolic machine learning system, or any combination thereof.

35. (Currently Amended) The system according to claim 25 wherein the means for ~~pattern analyzer~~ deriving the set of associations comprises an Adaptive Resonance Associative Map for analyzing the plurality of training samples.

36. (Currently Amended) The system according to claim 25 wherein the discovered knowledge comprises implicit hidden key relations between the sub-concepts ~~attributes of the entities~~ and the key concepts ~~entities~~.

37. (Currently Amended) The system according to claim 25 wherein the ~~knowledge discovery~~ system further comprises a user interface for displaying the semi-structured meta-data, the key concepts ~~entities~~, the key relations, the sub-concepts ~~attributes~~, and the knowledge discovered.

38. (Currently Amended) The system according to claim 25 wherein the knowledge discovery system further comprises a user interface for obtaining user's instruction for the key concepts ~~entities~~ and the key relations.

39. (New) The method of claim 1 wherein filtering the semi-structured meta-data includes automatically performing a statistical analysis upon the semi-structured meta-data.

40. (New) The method of claim 1 wherein the domain knowledge base is dependent upon one of a taxonomy, an ontology, a database, and a concept hierarchy network.

41. (New) The method of claim 1 wherein analyzing the plurality of training samples comprises mapping a plurality of sub-concepts to a key concept.

42. (New) The method of claim 1 further comprising deriving a set of IF-THEN rules corresponding to the set of associations, the set of IF-THEN rules corresponding to new symbolic knowledge.